

LAPORAN PENELITIAN KELOMPOK

Aplikasi Data Mining untuk Klasifikasi Kompleksitas Part Di Industri Manufaktur



Oleh :

Dra. Ernawati. M.T.
Parama Kartika Dewa, S.T., M.T.

FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ATMA JAYA YOGYAKARTA
YOGYAKARTA
2010

LEMBAR PENGESAHAN LAPORAN PENELITIAN

No. Laporan : _____

1. a) Judul Penelitian : Aplikasi Data Mining untuk Klasifikasi Kompleksitas Part di Industri Manufaktur.
b) Macam Penelitian : Laboratorium
2. Personalia Ketua Peneliti
a) Nama : Dra. Ernawati,M.T.
b) Jenis Kelamin : Perempuan
c) Usia saat pengajuan proposal : 41 Tahun 5 bulan
d) Jabatan akademik, golongan : Lektor Kepala, IV/a
e) Bidang peminatan : Soft Computing
f) Fakultas, Program Studi : Teknologi Industri, Teknik Informatika
3. Personalia Anggota Peneliti
a) Nama anggota : Parama Kartika Dewa, S.T.,M.T.
4. Jangka waktu penelitian : 6 bulan
5. Biaya yang disetujui : Rp. 3.740.000

Yogyakarta, 16 September 2010

Ketua Peneliti,

Anggota Peneliti 1,

Dra. Ernawati, M.T.

Parama Kartika Dewa, ST, MT

Mengetahui,

Kepala Laboratorium Komputasi FTI

Patricia Ardanari, S.Si., M.T.

Dekan Fakultas Teknologi Industri UAJY,

Ketua LPPM UAJY,

Ir. B. Kristyanto, M. Eng.,Ph.D.

Dr. MF. Shellyana Junaedi, SE.,M.Si.

INTISARI

Permasalahan pada perusahaan manufaktur adalah bagaimana melakukan prediksi waktu penyelesaian order dari konsumen dapat dilakukan dengan baik, sehingga kepuasan konsumen dapat dijaga dengan baik. Beragamnya kompleksitas order membuat prediksi waktu penyelesaian sulit untuk dilakukan. Kompleksitas yang dimaksudkan dalam penelitian ini dinyatakan dalam size dan ketebalan part.

Mencari perumusan proses prediksi waktu penyelesaian dengan menggunakan data mining adalah proses ekstraksi pengetahuan dari sekelompok data yang besar. Salah satu teknik dalam data mining adalah klasifikasi. Klasifikasi dilakukan untuk mengelompokkan sekumpulan data ke dalam kelas-kelas yang telah ditentukan terlebih dahulu berdasarkan kesamaan karakteristik yang dimiliki. Aplikasi data mining dengan menggunakan pohon keputusan diterapkan dalam penetapan perumusan prediksi waktu penyelesaian.

Hasil yang diperoleh berupa perumusan aplikasi pohon keputusan untuk melakukan prediksi waktu penyelesaian order dengan kelompok waktu penyelesaian cepat, sedang dan lambat.

KATA PENGANTAR

Puji syukur kami sampaikan kepada Tuhan Yang Maha Esa atas Berkat dan RahmatNya yang telah dianugerahkan kepada kami, sehingga penelitian dengan judul ” Aplikasi Data Mining untuk Klasifikasi Kompleksitas Part di Industri Manufaktur” dapat selesai.

Dalam menyelesaikan penelitian ini kami telah dibantu oleh pihak-pihak yang menaruh perhatian atas pentingnya penelitian ini. Kami ingin mengucapkan terima kasih kepada :

1. Bapak Ir. A. Koesmargono, M.Const.Mgt.Ph.D, selaku Rektor Universitas Atma Jaya Yogyakarta.
2. Ibu Dr. MF. Shellyana Junaedi, SE.,M.Si , selaku Ketua LPPM UAJY.
3. Bapak Ir. B. Kristyanto, M. Eng.,Ph.D, selaku Dekan Fakultas Teknologi Industri UAJY.
4. Ibu Patricia Ardanari S.Si., M.T., selaku Kepala Laboratorium Komputasi FTI.
5. Rekan-rekan asisten dan laboran yang telah membantu kami.

Demikian laporan yang kami sampaikan, semoga laporan ini berguna bagi komunitas di Universitas Atma Jaya Yogyakarta.

Yogyakarta, September 2010

Tim Peneliti

DAFTAR ISI

Halaman Judul	
Halaman Pengesahan	i
Intisari	ii
Kata Pengantar	iii
Daftar Isi	iv
Daftar Tabel	v
Daftar Gambar	vi
 BAB 1. PENDAHULUAN	
1.1. Latar Belakang	1
 BAB II. TINJAUAN PUSTAKA dan LANDASAN TEORI	
2.1. Tinjauan Pustaka	3
2.2. Data Mining	4
2.3. Klasifikasi	5
2.4. Pohon Keputusan	5
2.5. Information Gain	7
2.6. Part	8
 BAB III. PERUMUSAN MASALAH, TUJUAN PENELITIAN DAN MANFAAT PENELITIAN	
3.1. Perumusan Masalah.....	10
3.2. Tujuan Penelitian	10
3.3. Manfaat Penelitian	10
3.4. Batasan Masalah	10
 BAB IV. METODOLOGI PENELITIAN	
4.1. Jenis dan Lokasi Penelitian	12
4.2. Metode Pengambilan Data	12
4.3. Metode Analisis Data	12
 BAB V. DATA, ANALISIS, DAN PEMBAHASAN	
5.1. Data Penelitian	13
5.2. Analisis Data	16
5.3. Pembahasan	23
 BAB VI. KESIMPULAN DAN SARAN	25
 DAFTAR PUSTAKA	26

DAFTAR TABEL

Tabel 5.1 Data Penelitian	13
Tabel 5.2 Data Training	15
Tabel 5.3 Data Training dengan size = a	19
Tabel 5.4 Data Training dengan size = b	20
Tabel 5.5 Data Training dengan size = c	21
Tabel 5.6 Data Pengujian Model	22

DAFTAR GAMBAR

Gambar 2.1 Ketebalan suatu part.....	9
Gambar 2.2 Size suatu part.....	9
Gambar 5.1 Simpul Akar 01.....	17
Gambar 5.2 Simpul Akar 02.....	18
Gambar 5.3 Simpul Akar 03.....	20
Gambar 5.4 Pohon Keputusan Klasifikasi Waktu Perakitan part.....	21

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data Mining adalah proses ekstraksi pengetahuan dari sekelompok data yang besar. Salah satu teknik dalam data mining adalah klasifikasi. Klasifikasi dilakukan untuk mengelompokkan sekumpulan data ke dalam kelas-kelas yang telah ditentukan terlebih dahulu berdasarkan kesamaan karakteristik yang dimiliki. Klasifikasi banyak digunakan diberbagai bidang, misal aplikasi klasifikasi di bidang pengenalan pola, diagnosis medis, deteksi kegagalan di industri dan lain sebagainya.

Liao, T.W dkk (1999) mencoba memberikan perspektif tentang aplikasi data mining dalam kajian teknik industri. Bidang kajian yang diaplikasikan adalah bagaimana mengatur aktivitas proses pembuatan part yang berbasis pada teknologi proses “ Group Teknologi”. Teknologi ini mengandalkan kemampuan mesin yang dipergunakan atau dengan kata lain mengandalkan teknologi permesinaan yang dimiliki.

Karakter proses pembuatan part pada industri menengah di Indonesia masih mengandalkan tenaga manusia, seperti kebijakan yang dicanangkan oleh pemerintah. Industri yang didukung perkembangannya di Indonesia harus memiliki karakter padat karya (faktor manusia memiliki kontribusi terbesar dalam operasional industri). Permasalahan mendasar bagi industri yang berbasis produksi berdasar order adalah kemampuan untuk memprediksi waktu penyelesaian order tersebut. Masalah ini semakin kompleks bila karakter order pada perusahaan tersebut tidak berulang. Sehingga kemampuan untuk memprediksi waktu yang diperlukan untuk menyelesaikan order tersebut sangat diperlukan. Dengan aplikasi yang telah diperkenalkan oleh Liao, T.W dkk tersebut, kajian data mining sangat dimungkinkan untuk diterapkan pada pencarian solusi untuk melakukan prediksi waktu penyelesaian berdasarkan order

yang masuk. Order yang masuk dapat berupa part dengan berbagai kompleksitas part. Dalam hal ini kompleksitas part dinyatakan dalam size dan ketebalan part.

Variasi pada part berupa berbagai ragam size dan ketebalan, dapat dijadikan sebagai dasar untuk membuat klasifikasi dalam data mining. Proses selanjutnya adalah mengolah data tersebut sehingga dapat dijadikan sebagai bahan pendukung keputusan penetapan waktu yang diperlukan untuk menyelesaikan order tersebut.



BAB II

TINJAUAN PUSTAKA dan LANDASAN TEORI

2.1. Tinjauan Pustaka

Kristyanto B dan Dewa PK (1999), melakukan penelitian untuk mereduksi waktu perakitan dengan menggunakan metode rancangan perakitan (*Design For Assembly / DFA*) yang dikembangkan oleh Boothroyd dan Dewhurst. Perbaikan yang dihasilkan dari penelitian ini adalah sistem perakitan steker menghasilkan pengurangan komponen yang semula 5 komponen menjadi 2 komponen dengan perbaikan efisiensi perancangan dari 0,36 menjadi 0,79.

Pada penelitian yang dilakukan oleh Aldana W.A. (2000), telah dikembangkan model data mining yang diterapkan pada sistem informasi ERP dan ASP pada lingkup industri. Hasil yang diperoleh adalah sistem dengan menggunakan kombinasi ERP dan ASP untuk meningkatkan aktivitas manajemen data base dalam lingkup industri dengan lebih baik.

Penelitian yang dilakukan oleh Luthfi, E.T. (2009) mengembangkan model penggunaan data mining untuk melihat pola penjualan barang yang selanjutnya dapat dipergunakan untuk menyusun strategi penjualan baru. Algoritma data mining yang dipergunakan dalam penelitian ini adalah algoritma asosiasi. Hasil yang diperoleh adalah keterkaitan atau pola kemunculan barang dalam transaksi penjualan, yang pada akhirnya dapat digunakan untuk menyusun strategi penjualan.

Penelitian yang sekarang direncanakan ditujukan untuk mengaplikasikan model data mining pada proses estimasi proses pembuatan part pada industri manufaktur. Proses ini bermanfaat bagi pihak perusahaan untuk mengestimasi waktu pembuatan sebuah produk pada suatu order tertentu.

2.2. Data Mining

Data yang terdapat dalam perusahaan selalu bertambah berbanding lurus dengan aktivitas yang terjadi dalam perusahaan tersebut. Data yang berkaitan dengan proses pembuatan produk tentunya akan selalu bertambah berbanding lurus dengan semakin banyaknya jumlah konsumen yang melakukan pemesanan dan jumlah variasi produk yang dipesan oleh konsumen. Keadaan ini menimbulkan kondisi jumlah data yang tersedia dalam perusahaan berupa database semakin besar ukurannya. Pemanfaatan ketersediaan database yang terus membesar ini kurang diperhatikan dalam proses pembuatan keputusan dan seringkali yang terjadi adalah intuisi dari para pembuat keputusan ini yang mendominasi proses pengambilan keputusan ini. Berawal dari hal tersebut maka berkembanglah konsep data mining. Data mining adalah suatu bagian yang tak terpisahkan dari *knowledge discovery in databases* (KDD), yaitu suatu proses mengubah data mentah menjadi informasi yang berguna (Tan, Steinbach dan Kumar, 2006). Data mining mampu menganalisa data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pendukung keputusan. Hasil dari aplikasi data mining dievaluasi untuk menemukan suatu informasi/pengetahuan baru yang menarik dan bernilai bagi perusahaan, dan kemudian divisualisasikan agar mempermudah pengguna memilih informasi-informasi yang mempunyai arti guna mendukung pembuatan keputusan. Proses KDD secara garis besar meliputi tahap-tahap sebagai berikut :

1. Seleksi data

Seleksi data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *knowledge discovery in databases* (KDD) dimulai.

2. Pemrosesan awal/pembersihan data

Sebelum proses data mining dilaksanakan, perlu dilakukan proses pembersihan. Proses pembersihan data meliputi antara lain membuang duplikasi data, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan yang mungkin terjadi pada data.

3. Transformasi data

Proses transformasi pada data perlu dilakukan agar data sesuai untuk proses data mining.

4. Interpretasi/evaluasi

Setelah proses data mining dilakukan, akan didapatkan pola informasi yang perlu ditampilkan dalam bentuk yang mudah dimengerti dan diinterpretasikan oleh pihak yang berkepentingan. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

Secara umum, fungsi dalam data mining dapat dikelompokkan menjadi 2, yaitu prediksi dan deskripsi. Prediksi adalah penggunaan beberapa variabel untuk memprediksi nilai yang tidak diketahui atau nilai di masa yang akan datang dari variabel lain. Deskripsi adalah pencarian pola yang dapat diinterpretasikan manusia yang mendeskripsikan data yang ada. Yang termasuk dalam prediksi antara lain klasifikasi dan regresi, sedangkan yang termasuk deskripsi adalah klastering dan analisis asosiasi.

2.3. Klasifikasi

Klasifikasi adalah proses menemukan kumpulan pola atau fungsi-fungsi yang mendeskripsikan dan memisahkan kelas data satu dengan lainnya, dapat digunakan untuk memprediksi data yang belum memiliki kelas data tertentu (Han dan Kamber, 2001). Klasifikasi digunakan untuk memprediksi label kelas yang bersifat kategori.

Pada dasarnya ada dua langkah proses dalam klasifikasi data, yaitu membangun model berdasarkan sekumpulan data yang telah ditetapkan sebagai training set dan menggunakan model untuk klasifikasi. Banyak algoritma yang telah dikembangkan untuk melakukan proses klasifikasi, antara lain *Decision Tree* (Pohon Keputusan), Klasifikasi Bayesian, *Neural Network* dan lain-lain.

2.4. Pohon Keputusan

Pohon keputusan adalah salah satu dari teknik analisis yang digunakan untuk membuat keputusan. Pohon keputusan pertama kali diperkenalkan pada tahun 1960an oleh Fredkin. Pohon keputusan adalah sebuah struktur pohon mirip *flow chart*, dimana setiap simpul internal pada pohon merepresentasikan atribut

yang diuji, setiap cabang merupakan suatu pembagian hasil uji, dan simpul daun merepresentasikan kelompok kelas tertentu. Simpul akar dari sebuah pohon keputusan biasanya berupa atribut yang paling memiliki pengaruh pada suatu kelas tertentu. Pada umumnya pohon keputusan menggunakan strategi pencarian secara top-down untuk solusinya. Pada proses mengklasifikasi data yang tidak diketahui, nilai atribut akan diuji dengan cara melacak jalur dari simpul akar sampai simpul daun dan kemudian akan diprediksi kelas yang dimiliki oleh suatu data baru tertentu.

Berikut ini adalah algoritma pohon keputusan yang paling dasar (Han dan Kamber 2001) :

Algoritma : Generate pohon keputusan

Input : sampel training

Output : sebuah pohon keputusan

Metode :

- (1) Bentuk suatu simpul N;
- (2) Jika semua sampel berada pada kelas yang sama, C, maka return N sebagai simpul daun dengan label kelas C;
- (3) Jika list atribut kosong maka
return N sebagai simpul daun dengan label kelas yang paling sering muncul dalam sampel
- (4) Pilih atribut test, atribut dengan information gain terbesar;
- (5) Beri label simpul N dengan atribut test;
- (6) Untuk setiap nilai a_i yang diketahui dari atribut test
Bentuk cabang dari node N untuk kondisi atribut test = a_i
- (7) Misal s_i = kumpulan sampel dengan atribut test = a_i
Jika s_i kosong maka

Beri label pada simpul daun dengan kelas dari sampel yang paling sering muncul

(8) Else kembali ke langkah 1 dengan training sampel s_i

Pembentukan pohon dimulai dari simpul akar yang merepresentasikan semua data. Sesudah simpul akar terbentuk, maka data pada simpul akar akan diukur dengan information gain untuk memilih atribut mana yang akan dijadikan atribut test. Sebuah cabang dibentuk dari atribut test yang dipilih menjadi pembagi dan data akan didistribusikan ke dalam cabang masing-masing. Algoritma ini bersifat rekursif untuk dapat membentuk sebuah pohon keputusan. Ketika sebuah atribut telah dipilih menjadi simpul pembagi atau cabang, maka atribut tersebut tidak diikuti lagi dalam penghitungan nilai information gain. Proses pembagian rekursif akan berhenti jika salah satu dari kondisi dibawah ini terpenuhi:

1. Semua data dari anak cabang telah termasuk dalam kelas yang sama.
2. Semua atribut telah dipakai, tetapi masih tersisa data dalam kelas yang berbeda. Untuk kasus seperti ini, diambil data yang mewakili kelas yang terbanyak untuk menjadi label kelas pada simpul daun.
3. Tidak terdapat data pada anak cabang yang baru. Untuk kasus ini, simpul daun akan dipilih pada cabang sebelumnya dan diambil data yang mewakili kelas terbanyak untuk dijadikan label kelas.

2.5. Information Gain

Information gain adalah suatu ukuran yang digunakan untuk memilih atribut test pada setiap simpul dalam pohon keputusan. Atribut yang memiliki nilai information gain terbesar dipilih sebagai atribut test untuk simpul tersebut.

Misal S adalah himpunan yang memuat s sampel data. Misal ada m nilai berbeda yang mendefinisikan m kelas yang berbeda, C_i ($i = 1, 2, \dots, m$). Misal s_i jumlah sampel dalam C_i , maka informasi harapan yang diperlukan untuk mengklasifikasi suatu sampel yang diberikan adalah

$$I(s_1, s_2, \dots, s_m) = - \sum_{i=1}^m p_i \log_2(p_i) \quad \dots\dots\dots \text{(persamaan 2.1)}$$

dengan p_i = probabilitas suatu sampel sembarang masuk dalam kelas $C_i = s_i/S$.

Misal atribut A memiliki v nilai berbeda $\{a_1, a_2, \dots, a_v\}$. Atribut A dapat digunakan untuk membagi S menjadi v subset $\{S_1, S_2, \dots, S_v\}$ dimana S_j memuat sampel-sampel dalam S yang memiliki nilai a_j dari A. Jika a dipilih sebagai atribut test, maka subset ini akan berhubungan dengan cabang yang tumbuh dari simpul yang memuat S . Misal s_{ij} = jumlah sampel dari kelas C_i dalam subset S_j , maka informasi harapan berdasar pada partisi kedalam subset-subset oleh A dinyatakan sebagai

$$E(A) = \sum_{j=1}^v \frac{s_{1j} + \dots + s_{mj}}{s} I(s_{1j}, \dots, s_{mj}) \quad \dots\dots\dots \text{(persamaan 2.2)}$$

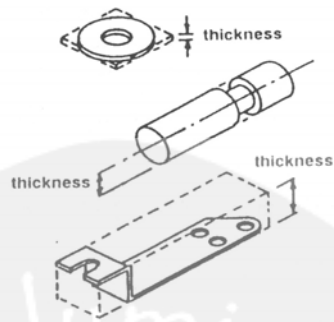
Sehingga Information gain yang didapat dengan mencabangkan pada atribut A adalah

$$\text{Gain}(A) = I(s_1, s_2, \dots, s_m) - E(A) \quad \dots\dots\dots \text{(persamaan 2.3)}$$

2.6. Part.

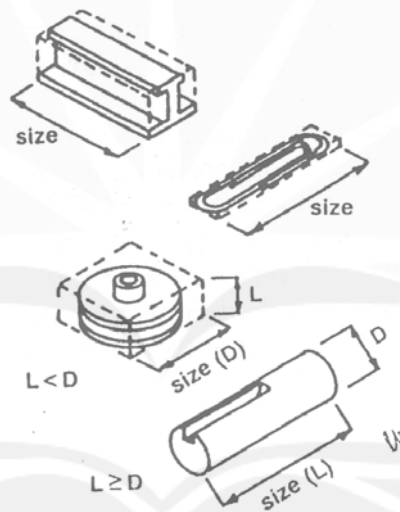
Part yang dimaksudkan dalam penelitian adalah bagian dari produk. Penetapan kompleksitas part yang dipergunakan dalam penelitian ini mengacu pada konsep yang dikembangkan oleh Boothroyd/Dewhurst (1996). Konsep ini memberikan beberapa kriteria pada sebuah part. Kriteria tersebut antara lain size dan ketebalan. Prinsip dari kriteria tersebut adalah efek dari size dan ketebalan, semakin kecil size dan ketebalannya maka akan semakin meningkat kesulitan untuk menangani part tersebut. Konsekuensi yang dihasilkan adalah waktu yang diperlukan untuk melakukan perakitan akan semakin lama.

Definisi ketebalan adalah dimensi panjang dari sisi terpendek bentuk part. Deskripsi pengertian dari definisi ketebalan dapat dilihat pada gambar 2.1 berikut.



Gambar 2.1 Ketebalan suatu part

Definisi size dari suatu part adalah panjang dari sisi dimensi terpanjang part. Deskripsi pengertian dari definisi size dapat dilihat pada gambar 2.2 berikut.



Gambar 2.2 Size suatu part

BAB III

PERUMUSAN MASALAH, TUJUAN PENELITIAN DAN MANFAAT PENELITIAN

3.1. Perumusan Masalah

Bagaimana melakukan prediksi waktu penyelesaian perakitan dengan menggunakan teknik klasifikasi part berdasarkan kompleksitas part yang dinyatakan dalam size dan ketebalan part.

3.2. Tujuan Penelitian

Mencari perumusan proses prediksi waktu penyelesaian waktu perakitan dengan menggunakan teknik klasifikasi part berdasarkan kompleksitas part yang dinyatakan dalam size dan ketebalan part.

3.3. Manfaat Penelitian

Hasil penelitian ini akan memberikan kontribusi pemanfaatan database yang ada di industri manufaktur sebagai dasar untuk prediksi waktu penyelesaian waktu perakitan dengan menggunakan teknik klasifikasi part berdasarkan size dan ketebalan part. Pada bidang kajian ergonomi dan sistem produksi akan diperkenalkan aplikasi data mining sebagai bahan penelitian lebih lanjut untuk mengembangkan sebuah sistem pendukung keputusan berkaitan dengan perancangan produk dan manajemen rantai produksi (*shopfloor*).

3.4. Batasan Masalah

Adapun batasan terhadap permasalahan yang diambil adalah sebagai berikut :

- a. Kompleksitas part dinyatakan dalam size dan ketebalan part.

- b. Variabel yang digunakan untuk klasifikasi adalah ketebalan dan size, sedangkan variabel hasil klasifikasi adalah waktu penyelesaian perakitan. Ketebalan dibagi menjadi 2 yaitu ≤ 2 mm dan >2 mm. Size dibagi menjadi 3 yaitu size ≤ 6 mm, size part lebih besar dari 6 mm sampai 15 mm, size > 15 mm. Waktu penyelesaian perakitan dibagi dalam 3 kelas yaitu lambat, sedang dan cepat.
- b. Teknik klasifikasi yang akan digunakan adalah decision tree.



BAB IV

METODOLOGI PENELITIAN

4.1. Jenis dan Lokasi Penelitian

Penelitian yang dilakukan adalah penelitian terapan yang berkaitan dengan aplikasi data mining di bidang manufaktur. Penelitian dilakukan dengan mengambil data perakitan part yang dilakukan di laboratorium Analisis dan Perancangan Kerja FTI-UAJY.

Kegiatan perakitan part yang dilakukan menggunakan part dengan variasi size dan ketebalan tertentu. Variasi ini diharapkan mampu merepresentasikan keragaman part pada suatu industri manufaktur.

4.2. Metode Pengambilan Data

Data dalam penelitian ini diambil dari hasil amatan dalam kegiatan perakitan yang dilakukan pada laboratorium Analisis dan Perancangan Kerja FTI-UAJY.

Proses pengambilan data dilakukan dengan menggunakan teknik jam henti, dimana operator melakukan aktivitasnya dan waktu yang diperlukan untuk melakukan aktivitas tersebut dicatat oleh operator dengan menggunakan stopwatch.

4.3. Metode Analisis Data

Data dianalisis menggunakan salah satu algoritma klasifikasi yaitu algoritma Pohon Keputusan. Dalam algoritma ini setiap alternatif kemungkinan yang akan terjadi ditelaah lebih lanjut.

BAB V

DATA, ANALISIS, DAN PEMBAHASAN

5.1. Data Penelitian

Data dalam penelitian ini diambil dari hasil percobaan dalam kegiatan perakitan yang dilakukan dalam laboratorium Analisis dan Perancangan Kerja FTI-UAJY. Part yang dipergunakan dalam penelitian ini dikelompokkan berdasarkan kriteria berikut :

- a. Ketebalan part dibagi menjadi dua bagian yaitu dikodekan dengan angka 1 dan 2. Kode 1 diperuntukkan untuk ketebalan part ≤ 2 mm dan kode 2 diperuntukkan untuk ketebalan >2 mm
- b. Size part dibagi menjadi tiga bagian yaitu dikodekan dengan huruf a, b dan c. Kode a diperuntukkan untuk size part ≤ 6 mm, kode b diperuntukkan untuk size part lebih besar dari 6 mm sampai 15 mm dan kode c diperuntukkan untuk size > 15 mm.

Waktu yang diperlukan untuk melakukan perakitan part dalam penelitian ini dinyatakan dalam satuan waktu detik. Data hasil amatan yang akan dipergunakan sebagai training set dalam penelitian disajikan dalam tabel 5.1 berikut ini.

Tabel 5.1 Data Penelitian

No	Ketebalan(1,2)	Size (a,b,c)	Waktu (detik)
1	2	c	1.13
2	2	c	1.5
3	2	c	1.8
4	2	c	1.2
5	2	c	1.7
6	2	c	1.6
7	2	b	1.43
8	2	b	1.8

Tabel 5.1 Lanjutan

No	Ketebalan(1,2)	Size (a,b,c)	Waktu (detik)
9	2	b	1.9
10	2	b	2.1
11	2	b	2.2
12	2	b	1.9
13	2	a	1.88
14	2	a	2.25
15	2	a	2.3
16	2	a	2.55
17	2	a	2.7
18	2	a	2.6
19	1	b	1.69
20	1	b	2.06
21	1	b	2.36
22	1	b	2.51
23	1	b	2.3
24	1	b	2.3
25	1	a	2.18
26	1	a	2.55
27	1	a	2.85
28	1	a	2.9
29	1	a	3
30	1	a	2.63

Sebelum dilakukan proses data mining, perlu dilakukan proses transformasi data. Variabel waktu penyelesaian perakitan yang merupakan data numerik, harus ditransformasikan menjadi data kategori. Waktu penyelesaian perakitan dikelompokkan menjadi 3 kategori, yaitu lambat, sedang dan cepat. Transformasi dilakukan dengan mengelompokkan waktu penyelesaian perakitan dibawah dan sama dengan rata-rata minus 1 standar deviasi ($\leq \bar{x} - s$) sebagai lambat, waktu perakitan antara rata-rata minus 1 standar deviasi sampai dengan rata-rata plus 1 standar deviasi ($\bar{x} - s < x < \bar{x} + s$) sebagai sedang dan diatas dan sama dengan rata-

rata plus 1 standar deviasi ($\geq \bar{x} + s$) sebagai cepat. Dari data diatas didapat rata-rata waktu perakitan = 2,13 dengan standar deviasi $s = 0,50$, sehingga waktu perakitan $< 1,63$ masuk kategori cepat, waktu perakitan antara 1,63 dan 2,63 masuk kategori sedang dan waktu perakitan $\geq 2,63$ masuk kategori lambat. Setelah dilakukan transformasi terhadap data tersebut didapat data training yang siap untuk diklasifikasikan (tabel 5.2)

Tabel 5.2 Data Training

No	Ketebalan(1,2)	Size (a,b,c)	Waktu
1	2	c	Cepat
2	2	c	Cepat
3	2	c	Sedang
4	2	c	Cepat
5	2	c	Sedang
6	2	c	Cepat
7	2	b	Cepat
8	2	b	Sedang
9	2	b	Sedang
10	2	b	Sedang
11	2	b	Sedang
12	2	b	Sedang
13	2	a	Sedang
14	2	a	Sedang
15	2	a	Sedang
16	2	a	Sedang
17	2	a	Lambat
18	2	a	Sedang
19	1	b	Sedang
20	1	b	Sedang

Tabel 5.2 Lanjutan

No	Ketebalan(1,2)	Size (a,b,c)	Waktu
21	1	b	Sedang
22	1	b	Sedang
23	1	b	Sedang
24	1	b	Sedang
25	1	a	Sedang
26	1	a	Sedang
27	1	a	Lambat
28	1	a	Lambat
29	1	a	Lambat
30	1	a	Lambat

5.2. Analisis Data

Setelah data yang diperoleh sebagai hasil amatan disajikan dalam format tabel, maka akan dilakukan proses klasifikasi terhadap data dengan menggunakan pohon keputusan. Tahapan awal dalam proses pembuatan pohon keputusan adalah dengan membuat simpul akar. Simpul akar ini akan dinyatakan dengan simbol sebuah lingkaran. Pada sebuah simpul akar akan terdapat variabel yang disebut dengan variabel test. Untuk menentukan variabel test pada simpul akar digunakan information gain. Variabel dengan information gain terbesar akan dipilih sebagai variabel test. Pada penelitian ini terdapat dua (2) variabel yang digunakan sebagai variabel klasifikasi, yaitu size dan ketebalan, maka dihitung information gain untuk size ($\text{Gain}(\text{size})$) dan information gain untuk ketebalan ($\text{Gain}(\text{ketebalan})$). Nilai dari masing-masing variabel ini ditetapkan dengan menggunakan persamaan 2.1, persamaan 2.2 dan persamaan 2.3. Hasil dari perhitungan dengan menggunakan persamaan-persamaan tersebut adalah sebagai berikut :

$$I(cepat, sedang, lambat) = -\frac{5}{30} \log_2\left(\frac{5}{30}\right) - \frac{20}{30} \log_2\left(\frac{20}{30}\right) - \frac{5}{30} \log_2\left(\frac{5}{30}\right) = 2,54$$

Tahap perhitungan selanjutnya menentukan variabel E (size) dan E (ketebalan), prosesnya adalah sebagai berikut:

$$E(size) = \frac{12}{30} I(0,7,5) + \frac{12}{30} I(1,11,0) + \frac{6}{30} I(4,2,0) = 0,74$$

$$E(ketebalan) = \frac{12}{30} I(0,8,4) + \frac{18}{30} I(5,12,1) = 1,05$$

Tahap berikutnya adalah menentukan Gain(size) dan Gain(ketebalan), proses perhitungannya adalah sebagai berikut :

$$Gain(size) = 2,54 - 0,74 = 1,80$$

$$Gain(ketebalan) = 2,54 - 1,05 = 1,49$$

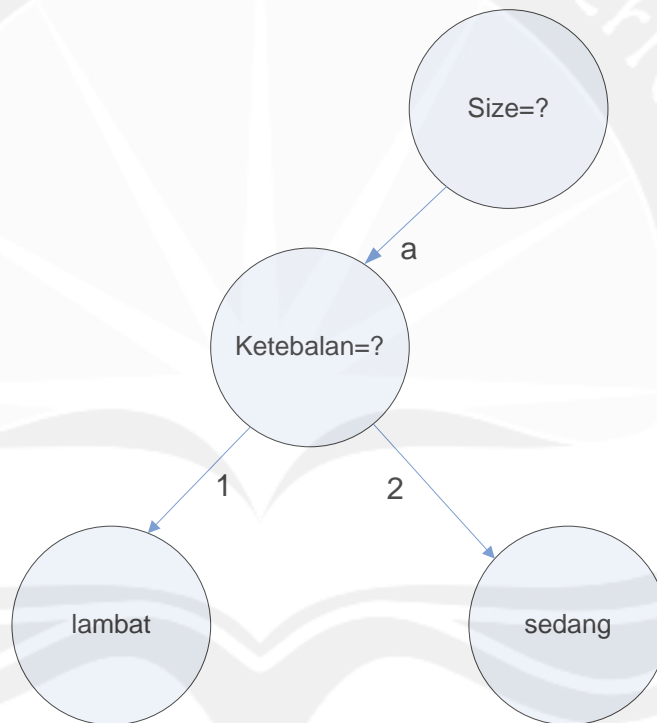
Berdasarkan hasil perhitungan diatas, karena $Gain(size) > Gain(ketebalan)$, maka simpul akar akan dipecah berdasarkan atribut size seperti pada gambar berikut :



Gambar 5.1. Simpul Akar 01

Dalam penelitian ini terdapat tiga nilai untuk atribut size yaitu a, b, dan c. Berdasarkan jumlah atribut tersebut, maka simpul yang terdapat pada gambar 5.1 dibuatkan cabang sebanyak 3 cabang. Ketiga cabang tersebut adalah cabang untuk size a, b dan c. Untuk cabang a, karena waktu perakitan tidak menunjuk pada suatu waktu perakitan tertentu, maka masih harus dicabangkan berdasarkan atribut yang tersisa, yaitu atribut ketebalan. Oleh karena itu dilakukan pengamatan

terhadap semua data training yang memiliki $size=a$. Dari table 5.3 tampak ada 12 data dengan $size=a$. Karena atribut ketebalan memiliki 2 nilai, yaitu ketebalan 1 dan 2, maka atribut ketebalan dipecah menjadi 2 cabang, yaitu cabang 1 untuk ketebalan 1 dan cabang 2 untuk ketebalan 2. Terdapat enam data dengan $size=a$ ketebalan 1, dimana empat diantaranya memiliki waktu perakitan lambat dan dua data lainnya memiliki waktu perakitan sedang. Berdasarkan sifat single majority maka dibentuk simpul daun dengan kelas waktu perakitan lambat. Untuk $size=a$ ketebalan 2 juga ada enam data, lima diantaranya memiliki waktu perakitan sedang, sehingga dibentuk simpul daun dengan kelas waktu perakitan sedang, seperti gambar berikut:



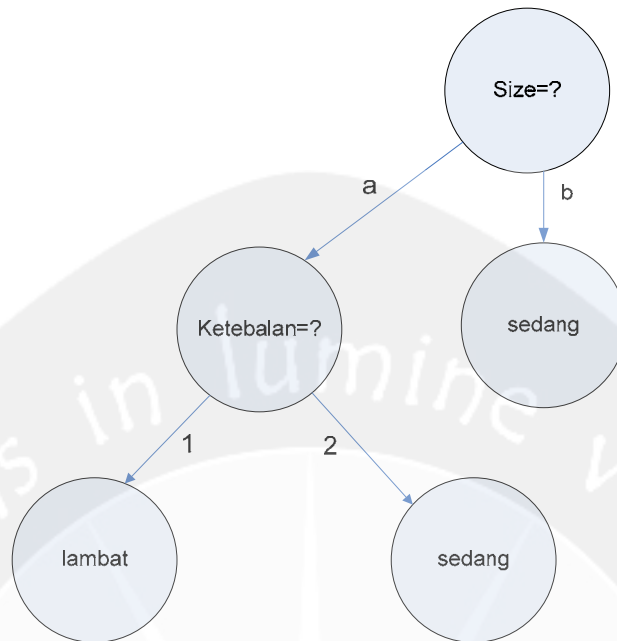
Gambar 5.2. Simpul Akar 02

Tabel 5.3 Data Training dengan size = a

No	Ketebalan(1,2)	Size (a,b,c)	Waktu
13	2	A	Sedang
14	2	A	Sedang
15	2	A	Sedang
16	2	A	Sedang
17	2	A	Lambat
18	2	A	Sedang
25	1	A	Sedang
26	1	A	Sedang
27	1	A	Lambat
28	1	A	Lambat
29	1	A	Lambat
30	1	A	Lambat

Tabel 5.3 merupakan data yang diambil dari data amatan dengan nomor urut pada tabel 5.3 tidak mengalami perubahan.

Untuk cabang b, yaitu data dengan size=b, ada 12 data (table 5.4). Seluruh data size b dengan ketebalan 1, memiliki waktu perakitan sedang, Sedangkan untuk size b dengan ketebalan 2, lima dari enam data juga memiliki waktu perakitan sedang, maka pada cabang b cukup dibentuk satu simpul daun dengan kelas waktu perakitan sedang, seperti gambar berikut :



Gambar 5.3. Simpul Akar 03

Tabel 5.4 Data Training dengan size=b

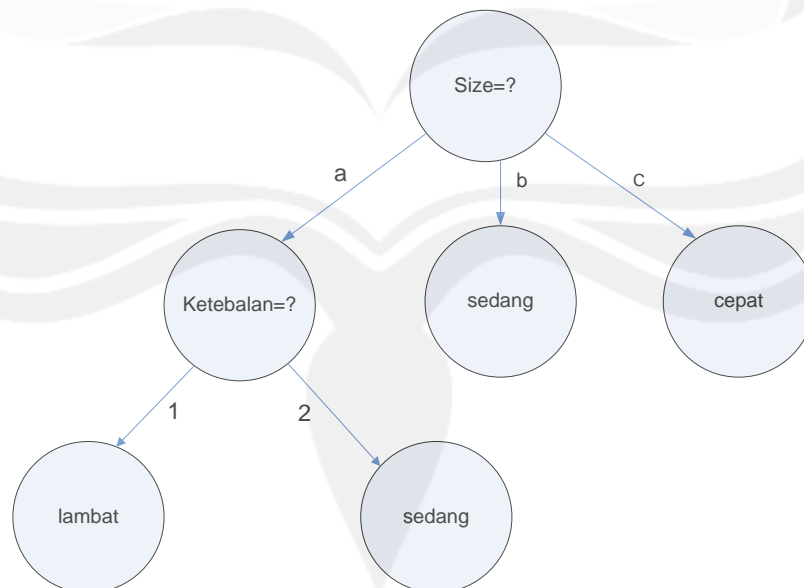
No	Ketebalan(1,2)	Size (a,b,c)	Waktu
7	2	b	Cepat
8	2	b	Sedang
9	2	b	Sedang
10	2	b	Sedang
11	2	b	Sedang
12	2	b	Sedang
19	1	b	Sedang
20	1	b	Sedang
21	1	b	Sedang
22	1	b	Sedang
23	1	b	Sedang
24	1	b	Sedang

Untuk cabang c, yaitu data dengan size=c, ada 6 data, semuanya memiliki ketebalan=2 (table 4.5). Karena mayoritas untuk size c waktu penyelesaian cepat, maka pada cabang c dibentuk simpul daun yaitu kelas waktu perakitan cepat.

Tabel 5.5 Data Training untuk size=c

No	Ketebalan(1,2)	Size (a,b,c)	Waktu
1	2	c	Cepat
2	2	c	Cepat
3	2	c	Sedang
4	2	c	Cepat
5	2	c	Sedang
6	2	c	Cepat

Pohon keputusan yang terbentuk berdasarkan seluruh alternatif yang mungkin terjadi adalah sebagai berikut :



Gambar 5.4. Pohon keputusan klasifikasi waktu perakitan part berdasarkan size dan ketebalan part

Dari pohon keputusan tersebut didapat 4 aturan untuk klasifikasi waktu perakitan part berdasarkan size dan ketebalan part sebagai berikut :

1. Jika size = a dan ketebalan = 1 maka waktu perakitan = lambat
2. Jika size = a dan ketebalan = 2 maka waktu perakitan = sedang
3. Jika size = b maka waktu perakitan = sedang
4. Jika size = c maka waktu perakitan = cepat

Sampai dengan tahap ini sebenarnya model yang diharapkan untuk merepresentasikan model estimasi waktu perakitan telah diperoleh. Model yang telah diperoleh ini perlu diuji terlebih dahulu untuk memastikan bahwa model yang diperoleh ini memang benar dapat dipergunakan. Tahapan berikut yang harus dilakukan adalah melakukan pengujian model.

Tahapan pengujian model dilakukan dengan mengambil data untuk pengujian. Untuk menguji model yang didapat dengan pohon keputusan tersebut, diambil 15 data sebagai data pengujian. Berdasarkan ketebalan dan sizenya dapat diprediksi waktu perakitan part dengan cara menerapkan aturan yang sesuai. Hasil pengujian dengan mengambil 15 data uji sebagai berikut :

Tabel 5.6 Data Pengujian Model

No	Ketebalan(1,2)	Size (a,b,c)	Waktu (dtk)	Waktu (sesudah transformasi)	Hasil Prediksi
1	2	C	1,3	cepat	Cepat
2	1	B	2,3	sedang	Sedang
3	2	C	1,8	sedang	Cepat
4	1	B	2,51	sedang	Sedang
5	2	A	2,3	sedang	Sedang
6	2	A	2,5	sedang	Sedang
7	1	A	2,18	sedang	Lambat
8	1	A	2,85	lambat	Lambat
9	2	B	1,8	sedang	Sedang
10	2	A	2,6	sedang	Sedang

Tabel 5.6 Lanjutan

No	Ketebalan(1,2)	Size (a,b,c)	Waktu (dtk)	Waktu (sesudah transformasi)	Hasil Prediksi
11	2	B	2,1	sedang	Sedang
12	2	B	1,9	sedang	Sedang
13	1	B	2,06	sedang	Sedang
14	1	A	2,57	Sedang	Lambat
15	2	C	1,5	cepat	Cepat

Dari hasil pengujian tersebut, didapat 3 hasil yang berbeda antara waktu perakitan part hasil prediksi dengan waktu perakitan part setelah transformasi. Berarti model yang didapat memiliki error sebesar 20%, atau memiliki ketepatan sebesar 80%. Hasil ini menunjukkan bahwa model yang telah diperoleh mampu untuk dipergunakan sebagai model untuk melakukan estimasi waktu perakitan dengan tingkat kesalahan yang mungkin terjadi sebesar 20%. Bila dipandang dari sudut pandang yang lain dapat dikatakan bahwa model untuk melakukan perkiraan waktu perakitan memiliki tingkat keakuratan sebesar 80%.

5.3. Pembahasan

Pembahasan dilakukan terhadap model untuk melakukan estimasi perakitan yang diperoleh dari hasil pengolahan data dan pemodelan yang telah dilakukan sebelumnya. Berdasarkan hasil klasifikasi yang dilakukan terhadap ketebalan dan size part untuk mendapatkan waktu perakitan part dengan mengambil 30 data training, didapat 4 aturan sebagai berikut :

1. Jika size ≤ 6 mm dan ketebalan ≤ 2 mm maka waktu perakitan dapat dikelompokkan dalam kelompok waktu lambat
2. Jika size ≤ 6 mm dan ketebalan > 2 mm maka waktu perakitan dapat dikelompokkan dalam kelompok waktu sedang
3. Jika size lebih besar dari 6 mm sampai 15 mm maka waktu perakitan dapat dikelompokkan dalam kelompok waktu sedang

4. Jika size > 15 mm maka waktu perakitan dapat dikelompokkan dalam kelompok waktu cepat

Setelah dilakukan pengujian terhadap model tersebut, dengan mengambil 15 data pengujian, ternyata keakuratan dari model sebesar 80%. Walaupun keakuratan sudah cukup baik, namun keakuratan model dapat lebih ditingkatkan dengan mengambil data training yang lebih besar.

Model yang didapat dari penelitian ini dapat digunakan di bidang industri manufaktur untuk mendukung keputusan dalam memprediksi waktu penyelesaian order. Hal yang perlu diperhatikan adalah dalam penetapan waktu penyelesaian dengan menggunakan model ini perlu dipertimbangkan untuk memberikan waktu kelonggaran yang diperuntukkan untuk mengantisipasi error sebesar 20%.

BAB V

KESIMPULAN DAN SARAN

Penelitian ini memberikan hasil bahwa dengan menggunakan konsep klasifikasi data mining yang diterapkan dalam aktivitas perakitan part dalam industri mampu memberikan kemampuan memprediksi waktu perakitan yang diperlukan bagi part dengan tingkat kompleksitas tertentu. Dalam penelitian ini kompleksitas yang dimaksud merupakan kombinasi antara ketebalan dengan size part yang berbeda. Aplikasi praktis yang dapat disumbangkan bagi kegiatan manufaktur adalah pihak manufaktur dapat melakukan prediksi lebih akurat berkaitan dengan waktu yang diperlukan untuk menyelesaikan suatu order yang datang. Ujung-ujung dari aplikasi praktis ini adalah meningkatkan kepuasan konsumen karena estimasi waktu penyelesaian suatu order dapat ditentukan dengan lebih tepat.

Saran pada penerapan hasil penelitian ini adalah diperlukan dukungan database berkaitan dengan waktu penyelesaian part yang dimiliki oleh perusahaan manufaktur. Semakin banyak data keragaman part dan waktu penyelesaiannya akan meningkatkan ketepatan dalam memprediksi waktu penyelesaian suatu order. Pengembangan penelitian ini dapat dilakukan dengan dikembangkannya perangkat lunak yang mampu mengolah database yang telah ada di perusahaan manufaktur. Diharapkan perangkat lunak ini dapat menjadi bagian dari sistem pendukung keputusan perusahaan dalam memberikan pelayanan kepada konsumen.

Daftar Pustaka

- Aldana. W.A., Data Mining Industry: Emerging Trends and New Opportunities, Thesis, Master of Engineering in Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, 2000
- Luthfi, E.T., Penerapan Data Mining Algoritma Asosiasi untuk meningkatkan penjualan, Jurnal DASI, vol. 10, No 1 Maret 2009.
- Giudici,P., Figini,S., Applied Data Mining for Business and Industry, 2 nd Edition, Wiley, 2009.
- Kristyanto, B dan Dewa, PK., Kontribusi Ergonomi Untuk Rancangan Perakitan, Jurnal Teknologi Industri, Vol III, No 1, ISSN 1410-5004, 1999
- Tan, Steinbach dan Kumar, 2006, Introduction to Data Mining, Pearson Education Inc.,
- Han dan Kamber, 2001, Data Mining Concepts and Techniques, Academic Press